DOCUMENT RESUME

ED 025 804 24 CG 003 462

By-Schoenfeldt, Lyle F.; Cooley, William W.

Data Archives as Resources for Research, Instruction, and Policy Planning: The Project Talent Data Bank.

American Institutes for Research, Pittsburgh, Pa.; Pittsburgh Univ., Pa.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No-BR-5-0606
Pub Date [68]
Contract-OEC-6-10-065
Note-25p.
EDRS Price MF-$0.25 HC-$1.35
Descriptors-*Computer Science, *Data Processing, Higher Education, Information Processing, *Information Storage, Instruction, Policy Formation, Research Tools, Secondary Education

This paper describes the advances in storing and accessing large resources of data as well as possibilities for utilizing these archives in research, instruction, and policy planning. Particular reference is made to the Project TALENT Data Bank. Project TALENT is described briefly. A section summarizes the state of computer technology with respect to the application of computers to data archives and higher education. Other data archives are described which are available to behavioral, social, and educational researchers. The advantages in using the holdings of a data bank are noted. The operations of the Project TALENT Data Bank are described in three areas: research, instruction, and the planning of social policy. (BP)

ED025803

# DATA ARCHIVES AS RESOURCES FOR RESEARCH, INSTRUCTION, AND POLICY PLANNING: THE PROJECT TALENT DATA BANK[1]

Lyle F. Schoenfeldt

American Institutes for Research

and

William W. Cooley

University of Pittsburgh

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

[1]Project TALENT is being carried out by the Amer. Inst. for Resch. and the Univ. of Pittsburgh pursuant to contract with the Office of Education, U.S. Department of Health, Education, and Welfare (Contract Number OE-6-10-065 Project Number 3051). The authors are grateful to Marsha D. Brown for assistance in preparation of this article.

EA 003 462

In the most general sense, a data archive is a library of data. As with libraries of books, holdings are carefully catalogued and cross-referenced in an attempt to provide ready access to the systematically organized collections of information. The necessity for maintaining libraries for written materials has long been recognized as critical for both research and educational processes. The notion of scientists benefiting from the previously recorded observations of other scientists also has a long and significant history, especially in the physical and biological sciences. In disciplines such as astronomy and geology, where experimentation is difficult if not impossible, systematic observations over time have been crucial to important theoretical developments.

The formation of data archives for behavioral, social, and educational research did not gain momentum until the late fifties and early sixties when the computer became available to a sizable portion of the research community. In fact, unlike libraries for published results, the repositories for storing and circulating the tremendous quantity of data constantly being gathered by behavioral science research workers are largely dependent on the unique capabilities of the computer.

The purpose of this paper is to describe the advances in storing and accessing large resources of data as well as the possibilities for utilizing these archives in research, instruction, and policy planning. Particular reference will be made to the Project TALENT Data Bank, a resource designed as potentially valuable to behavioral scientists with research concerns beyond those central to the goals of the project.

It is hoped that this description will have a twofold outcome: first, that it will stimulate scientists to organize the data they are collecting or planning to collect so they may be generally available for other research purposes. Secondly, this description should encourage scientists, instructors, and policy planners to consider data banks as viable resources for their particular needs.

The paper is divided into the following five sections. Since the value of the Project TALENT data as a resource depends in part on the sample tested and measurements obtained, a brief description of this project comprises the first section. Because behavioral science data archives have been heavily influenced by advances in computer technology, the second section summarizes the state of the art with respect to the application of computers to data archives and higher education. The third section describes other data archives available to behavioral, social, and educational researchers, and the fourth section describes the advantages and disadvantages in using the holdings of a data bank. The fifth and final section describes the operations of the Project TALENT Data Bank in each of three areas: research, instruction, and the planning of social policy.

## Project TALENT

In the spring of 1960, Project TALENT began collecting data on 440,000 United States high-school students. The sample was carefully drawn so that it would be representative of all public, private, and parochial school students in grades nine through twelve. The two-day battery

of tests administered in that first data collection effort consisted of 60 different ability tests, over 30 noncognitive measures, and about 400 items of personal history, family background, and educational and career plans.

In addition to the data obtain d from each student, information was collected from the 1,300 schools attended by these students. Personnel from the schools completed three questionnaires: one on the school characteristics, another on the guidance program, and a third on the specific counseling services in each school. Detailed descriptions of the Project TALENT sample and the original test battery can be found in Design for a Study of American Youth (Flanagan, et al., 1962). The American High-School Student (Flanagan, et al., 1964) reported the abilities, interests, plans, and activities of high-school students as determined by the 1960 testing.

Subsequent to the initial data collection in the spring of 1960, students were contacted through follow-up questionnaires one year after their class graduated from high school. In addition, data are currently being collected from these people five years out of high school so that it is possible to study not only college entrance or first jobs, but completion of college, entrance to graduate school, five-year work histories, marriage, etc. Table 1 summarizes the complete follow-up schedule, part of the original plan for Project TALENT. The Project TALENT One-Year Follow-Up Studies (Flanagan, et al., 1966) described the follow-up procedures (including correction for nonrespondent bias) the questionnaires used, and the relationships between follow-up

information and data from the same people in the 1960 testing. A recent publication by Cooley and Lohnes (1968) reports the initial analysis of relationships between five-year follow-up information and 1960 data.

In order to assess changes occurring over the last three years of high school, a sample of over 100 of the original high schools was contacted and their seniors retested in the spring of 1963. Assuming normal class progression, individuals originally tested as ninth graders in these high schools were seniors in 1963. These seniors were retested with the original 1960 tests and inventories. These data are described in The High School Years: Growth in Cognitive Skills (Shaycoft, 1967).

Additional files in the TALENT library include: (1) special samples that participated in the regular testing, (2) additional data available for subsamples, and (3) a special retest study (Carp & Schoenfeldt, 1968; Schoenfeldt, 1968).

## Computer Applications, Data Archives, and Higher Education: The State of the Art

The development of data archives for behavioral, social, and educational researchers represents a significant step in the evolution of research strategies, a step that can be attributed directly to the advent and recent widespread availability of the electronic computer. Since the progress to date and future prospects for the establishment and utilization of data archives are so intimately tied to computer technology, this section will review the state of the art with respect to the diffusion of computer resources. It will be shown that: (1)

computer technology should be available to all researchers within the next several years; (2) increasing emphasis will be placed on the value of incorporating computer applications into the higher education of all students, including those in the behavioral sciences; and (3) data archives are an ideal resource for providing a majority of behavioral science students with experience in the application of the computer to the retrieval and analysis of data as well as insights into the research process.

Project TALENT provides an example of how the computer has facilitated not only large-scale research, but also the sharing of data. When the goals of TALENT were set forth in the mid 50's, computers were still in their infancy and the concept of a computerized data bank was scarcely considered. The data collection phase of TALENT was initiated a mere 14 years following the completion of the first electronic computer, the ENIAC, and less than ten years following the commercial availability of what is now termed the first generation computers (Borko, 1962; Green, 1963). The rapid development of the computer field and the concept of data banks has made possible the sharing of these data among scholars with diverse interests throughout the country.

Several recent reports have examined the role of the computer in the educational context, and have reached the following conclusions (McCarthy, 1966; Science Advisory Committee, 1967). First, the availability of computer services has not been uniform across institutions of higher education nor across departments within institutions that have computer services (Kemeny & Kurtz, 1968; Science Advisory Committee, 1967).

This problem of inadequate computer services at many institutions of higher education, can be attributed to the expense of such facilities. Two recent developments offer promise toward solution of this problem. The first is the establishment of an Office of Computing Activities in the National Science Foundation with specific responsibility for upgrading facilities at colleges and universities. A second development offering promise to colleges with inadequate computer facilities concerns technological advances in time-sharing (Fano & Corbató, 1966; Kemeny & Kurtz, 1968). Using time-shared facilities it is possible for smaller colleges and universities to have direct links with nearby institutions. This centralization of resources should benefit all institutions by permitting better facilities than any single center could afford.

As might be expected, utilization of the computer by departments and divisions within universities has varied tremendously, with the physical sciences and engineering tending toward extensive use while the social sciences, education, and arts make only slight to moderate use of the facilities (Science Advisory Committee, 1967). In many respects this problem is more difficult to overcome. Applications of computer technology in the process of higher education for students in physical science and engineering are considerably more obvious than for students in liberal arts. This difference is magnified by the fact that the faculty in the former areas are more accustomed to incorporating computer excercises into the educational process than are faculty in the social sciences.

Analysis of these problems leads to the second major conclusion of the previously cited reports -- the need for all university and college students to have some basic understanding of digital computation.

To quote from the Science Advisory Committee (1968),

> "...we believe that the computer and computing are rapidly coming to have an impact on the life of practically every member of our society. Most people educated beyond the high school level will have occasion to make use of these tools, and all will need sufficient understanding of their possibilities and limitations realistically to appraise the new opportunities now available for information processing (p. 23)."

With respect to the social sciences and education, the goal should be twofold. First is the necessity for a division of labor which frees the student from the burden of statistical and arithmetical chores for the study of theory, design, and interpretation. The one area in which this idea has achieved substantial implementation is in the instruction of statistics. For example, the definitive text on factor analysis by Horst (1965) includes computer programs to assist the student with the computational burden. A more recent introductory text integrates the theory of statistics with the application of the computer to this subject (Lohnes & Cooley, 1968). With respect to other areas of instruction, it will be necessary to create ways of combining computer applications with substantive material to attain a new and higher level of pedagogy.

It is suggested that the ideal vehicle for reaching the next plateau of instructional sophistication will be specially prepared files of data from archival holdings, computer programs designed for their analysis, along with adequate documentation of both files and programs. The continued expansion of computer facilities at institutions of higher education now makes it possible for upper level undergraduates and graduate students to share the benefits of the data collected in conjunction with

research projects from throughout the country. The outline of a program for the application of the data collected in conjunction with Project TALENT to such an instructional undertaking is described later in this paper. The program's goal is to introduce a new potency to the instruction of courses offered by a number of social science and education departments which have more often than not been merely a sterile review of past studies.

## Major Data Archives

The increasing interest in libraries as resources for facilitating maintenance and circulation of data is best exemplified by the enumeration of the major archives in Table 2. Most of these archives have been established within the past ten years. The typical evolutionary progression is service to a small special population followed by emphasis on providing service to the scholarly community in general.

The results of the present survey described in Table 2 are actually the nucleus of what will soon be a much greater number of data banks, perhaps two or three times as many within the next decade. Numerous individuals and organizations are in various stages of planning for the future storage and dissemination of research data.

A number of similarities in operational procedures are characteristic of the major archives. Most have policy committees that provide guidelines concerning criteria governing use of materials as well as acquisition of new data. Materials are generally supplied on a cost-reimbursable basis in accordance with the computer and personnel time required to fill the request. A large number have technical facilities for preparing special analyses, although the costs are often high if the size of the

data base is large, as with the Bureau of the Census, Project TALENT, and others. Descriptive materials of specific holdings and policies governing their utilization can generally be obtained by writing to the director of the archive.

In recognition of the need for coordination of activities two major groups have evolved. The Council of Social Science Data Archives (CSSDA) was initially formed in 1962, and established in its present form three years later. It is supported by the National Science Foundation and currently consists of 27 member data banks. It has sponsored five national meetings to facilitate exchange of information among members and interested non-members, and maintains committees for establishing standards in many areas of data bank development. It also cooperates with the major computer manufacturers to facilitate the development of special hardware and software components needed to effectively solve the special retrieval problems characteristic of data archives. The organization and operations of the CSSDA have been described by Glaser and Bisco (1966) and Bisco (1967).

The International Social Science Council (ISSC) formed a Standing Committee on Data Archives in 1966 to advance international cooperation among archives. The Committee utilizes Social Science Information, the official journal of the ISSC, to provide systematic coverage of data bank activities.

## Advantages and Disadvantages of Data Archives

The most attractive advantage of utilizing existing data files is economy, both in terms of time and money. These economies can best be appreciated by consideration of two different types of archives: Project

TALENT and the Roper Public Opinion Research Center. Both organizations collect and maintain data bases consisting of numerous files of information. Procedures have been established for making these data bases available to a broad spectrum of users.

The Project TALENT Data Bank refers to all the information collected in 1960 plus the information accumulated in successive follow-up surveys. It includes school and community data as well as individual data. The feature which distinguishes it from most other archives is the fact that new subjects are not being added, but rather additional data is continually being collected on the 440,000 original participants. An investigator interested in studying a long-range process, such as career development or marital adjustment over a ten-year period, ordinarily has many years to wait between the collection of data and conclusions concerning the behavior he is studying. By using data from the TALENT files, or some similar source, he would benefit from several advantages. First would be a much larger and more representative sample than could probably be justified for a single purpose study. Second, the initial contact with the subjects would have been made while the criterion behavior (marriage, career, etc.) was only in the planning stages, and thus it would be possible to observe more accurately the dynamics of the decision process. A third advantage would be a wider range of cogntive and noncognitive measurements than could be collected to investigate a single behavior domain.

The Roper Center now has available the basic data from almost 8,000 pulic opinion surveys, and is acquiring between 350 and 450 new surveys each year. The studies date from 1936, and were placed at the Center

by over 100 cooperating research organizations located in 44 countries. Approximately 80 percent of the surveys contain questions on a variety of subject matters (public affairs, market research, etc.) with the remainder being devoted to a single subject.

The Roper Center and similar libraries permit many research possibilities that would otherwise be unavailable. For example, the archive facilitates reanalysis of already published statistics in accordance with new perspectives. Such an opportunity is of particular value to social and behavioral scientists examining attitudes or opinions over an extended period of time. Since many of the archives have data from several countries, international comparative analyses can be undertaken by interested scholars.

Still another possibility is the combination of material from several archives (or data collected by the investigator with data retrieved from an archive) to obtain additional refinement in the level of analysis. For example, regional or local employment data from the Manpower Information System of the Bureau of Labor Statistics can be combined with consumer behavior information from the Roper Center. The follow-back technique has become an increasingly popular application of the TALENT data (Schoenfeldt, 1968). This procedure capitalizes on the premise that approximately five percent of those born between 1942 and 1946 participated in the 1960 testing. The method allows the investigator both to overcome two shortcomings of the follow-up procedure likely to restrict the utility of the data, i.e. limited questionnaire

space and nonresponse, and to obtain extensive aptitude and achievement information on subjects for whom he has criterion data. For example, a university could retrieve the records of TALENT participants who enrolled in their particular institution, and combine college performance data with the TALENT scores. (Procedures employed to insure adherence to confidentiality agreements with the participants will be described later in this section.) Thus the follow-back procedure allows researchers to obtain a representative sample of a specific group, append the 1960 test record with detailed criterion information, and investigate predictor-criterion relationships.

Another advantage of using data collected by others is the reduction in biases which are sometimes characteristic of data collected by a single investigator to support a particular hypothesis. Also, a large single collection effort often uses resources unavailable to an investigator undertaking a small study for the purpose of eliminating biases such as nonresponse to mailed questionnaires, sampling error, etc.

The biggest and most urgent problem facing those collecting, maintaining, or using large files of data on individual people is that of preserving each individual's right to have the information treated with confidentiality. The need for security is compounded by the fact that much of the data is collected under legal compulsion (Bureau of the Census, Social Security, Internal Revenue Service, etc.) or involuntarily (credit files). This problem has received wide attention, much of it unfavorable, as a result of several reports to the U.S. Bureau of the

Budget recommending a National Data Center (Dunn, 1965; Kaysen, 1967). An issue of the American Psychologist (May 1967) was devoted to the topic of privacy in behavioral research and more recently an article by Sawyer and Schechter (1968) discussed the importance of privacy with respect to social science utilization of the National Data Center.

With one exception, data banks designed for research purposes have no need for detailed information that could link a record to a specific individual. The exception arises when it is necessary to merge data from several archives. In this case the identifying information can be immediately and automatically removed from the file as a part of the final check of the merge process. Thus the scientist can have the large pool of data needed for studies which require Census data linked with TALENT ability data, for example, without being able to identify individuals.

The fact that the scientist does not need identifying information as a part of the data he analyzes does not deny that the problem of potential abuse is a real one for those maintaining data banks. As computers increase in importance, as information processors and files tend to become more and more centralized, it is necessary to develop fail-safe procedures for insuring confidentiality. McCarthy (1966) has outlined a "bill of rights" which might curb existing abuses (such as widely available credit information being used for purposes irrelevant to credit decisions) as well as remove the threat of future abuse. Among other things, his "established rights" would provide for public rules concerning access to the files, the opportunity

for the individual to read his own file and impose certain restrictions on access, recording and reporting of all entries to a file.

An objection frequently raised against the data bank concept is that one seldom finds data in existing archives which are exactly what the investigator would have collected himself to answer a particular research question. Thus, it is argued that one would tend to study questions for which data are already in the archive, rather than seek the data for the important questions of concern to the investigator. There is no doubt that this is sometimes the case, but it is certainly an inadequate reason to completely abandon the data bank concept. Experience at TALENT suggests that more often than not the available data are appropriate for investigating hypotheses posed by outside scientists with many and varied interests. More often than not the research process is one of successive approximations. For this reason the qualified conclusions resulting from studies which require some compromises can be of substantial value.

Project TALENT as a Resource for Research, Instruction, and Policy Planning

Research utilization. The Project TALENT Data Bank, formed by agreement with the U.S. Office of Education, was officially launched in May 1965 to demonstrate that researchers not members of the TALENT staff could successfully access the TALENT data base. This goal has been fulfilled. In the intervening period over 80 research projects have been completed on a cost-reimbursable basis.

Basic policy and guidelines regarding data utilization have been formulated by a panel of scientists. This panel serves in an advisory capacity and develops additional policy as necessary. The utilization

guidelines are administered by a local committee of researchers. Each request for utilization of TALENT data is screened by the local committee to insure compliance with confidentiality agreements given to participants and other policy guidelines.

For the most popular type of data bank project a computer printout is the final product. The outside researcher designs the analyses to be undertaken and receives the computer work completed to his specifications in printout form. He then interprets the data and reports the results.

For a second form of utilization, the Project TALENT staff proceeds from designing the analysis to writing a report of the results. This type of project is done primarily for agencies or associations that are interested in a problem for which the TALENT data are relevant but who do not have their own research staff.

A third type of data bank project is the release of actual data to the interested scholar. In this case, of course, all identifying information is removed from the work tape in compliance with agreements to maintain confidentiality.

Instructional utilization. In the realization that the extensive TALENT data could be a valuable adjunct in the instruction of many upper level undergraduate and graduate courses, a program for the preparation and dissemination of specially prepared files is being undertaken. It is envisioned that this program will be a partnership between Project TALENT and academic institutions to form an inter-disciplinary and inter-university training and research facility. The goals of this program are as follows: (1) creation of multi-purpose files that will serve a variety of training and research needs;

(2) development of computer systems of data retrieval and analysis designed to maximize the utility of the data for instructional purposes; and (3) development of specific instructional programs shaped to enable scholars to readily incorporate the data files into their undergraduate and graduate courses.

The potential for such a program is almost unlimited. First, it can provide the vehicle necessary to incorporate computer applications into many areas of higher education. Second, and more important, it will provide current data for many areas of instruction which, because of the lag between the initiation of research and incorporation of substantive results into text books, tend to become rather sterile reviews of history. The files would be constructed to permit a wide range of both typical and innovative studies. A third benefit derived from a common set of files made available to scholars in many different areas will be the promotion of inter-disciplinary considerations in research and instruction. As an example, the files would have application in the following disciplines: (1) psychology, (2) sociology, (3) economics, (4) human behavior-genetics, (5) education, (6) political science, (7) geography, and (8) medicine. The number of specific courses for which the special TALENT files could be utilized would vary for each discipline, with psychology, education, and sociology having a large number and the remaining areas having a more limited range of application.

Policy planning utilization. The capability of a large-scale longitudinal data resource to have a continuing influence on educational policy is a major factor which continues to differentiate Project TALENT from other major testing efforts. For example, data collected in

conjunction with the Equality of Educational Opportunity testing headed by Coleman, et al., (1966) have been made available to scientists for secondary analysis. However, design for this survey was cross-sectional rather than longitudinal, a fact which will result in the rapid depreciation of the data within the next few years. The expensive National Assessment testing program now underway has the same cross-sectional design. Whether or not the data collected will eventually be made available for secondary analysis, the opportunity for them to have a continuing influence on educational policy will be severely limited by their timeliness.

The TALENT Data Bank recently initiated a special service in recognition of an obligation to provide data for persons responsible for making educational policy decisions, or responsible for advising those who formulate educational policy. Certain types of data requests can be answered rapidly--usually within 48 hours; the general characteristics of requests that qualify for this priority service include minimal computer and/or personnel operations in the following areas: (1) variable recoding, (2) sample selection, and (3) limited analyses.

In the future, Project TALENT hopes to expand this service to include a telecommunications system capable of receiving and transmitting information to a terminal remotely located. Thus an individual researcher with access to a terminal unit would receive immediate feedback of analytic results from specific inquiries to the data bank. A pilot experiment with such a system was recently undertaken when Project TALENT participated in a conference on "The Use of Archives for the Study of Public Policy Issues" held in Pittsburgh under the auspices of the Council of Social Science

Data Archives. TALENT data and programs were adapted for telecommunications use, designed to maximize output flexibility and minimize input complexity and turn-around time. With this set-up it was possible for conference participants to communicate directy with a segment of the TALENT data base located in Palo Alto. Initially, the participant submitted a small number of one-digit parameters requesting specific subsets of data and specific analyses; he immediately received information on whether the parameters were accepted as legal or rejected as illegal and for what reason. Then, if the parameters were accepted, he received displayed on the terminal unit, the analyses requested. Total interaction time for each request was only a few minutes. Although the capabilities of the file and program used in this demonstration were limited, they pointed to some of the many possibilities for such a telecommunications system, especially in the field of policy planning. It is envisioned that the files and programs prepared for dissemination by one institution could be made available to outside users at other institutions by means of a telecommunications hookup similar to that used for the TALENT demonstration. Thus, through the use of many archives such as Project TALENT, those responsible for forming local, state, and national policy could, in the words of Time (1968, p. 13) use "a computer console to tap the memory bank of social knowledge and data". It is toward this end that the Council of Social Science Data Archives now is coordinating an experimental attempt to link the computer data bases of large archives.

## Summary and Conclusions

This paper has described the advances in storing and accessing large data bases as well as the possibilities for utilizing these resources in research, instruction, and policy planning. The result of these advances in computer technology has been the formation of a large number of general and special purpose archives for the behavioral and social sciences and education. Characteristics of the major data banks and the two groups which have evolved to coordinate their activities are described.

Project TALENT, a large-scale longitudinal study involving the testing of over 400,000 students and periodic follow-up contact for a 20 year period, is cited as one example of the range of advantages made possible through the data bank concept. In brief these advantages are: (1) economy of time that would otherwise be required to collect data and await the maturation of criterion information; (2) economy in terms of expense; (3) a larger and more representative sample of respondents than could be obtained by any single investigator; and (4) reduction in biases sometimes characteristic of data collected by a single investigator.

The biggest and most urgent problem facing those maintaining or using large files of data on individual people is that of preserving confidentiality. It is concluded that although the researcher does not ordinarily need identifying information as a part of the data he analyzes, the problem of potential abuse cannot be minimized.

Specific examples of present and future applications of data bank holdings in the areas of research, instruction, and planning of social policy are described.

## Bibliography


Bisco, R. L. Social science data archives: Progress and prospects. Social Science Information, 1967, 6 (1), 39-74.

Borko, H. Computer applications in the behavioral sciences. Santa Monica, California: Systems Development Corporation, 1962.

Brown, Marsha D. The Project TALENT demonstration: An application of computer telecommunications in the social sciences. In W. W. Cooley (Chm.), Data banks and educational policy planning. Symposium presented at the Fifth Annual Conference of the Council of Social Science Data Archives, Pittsburgh, June 11, 1968.

Carp, A. & Schoenfeldt, L. F. Project TALENT: A computer based study. Journal of School Psychology, 1968, 6, 186-192.

Coleman, J. S., Campbell, E. Q., Hobson, Carol J., McPartland, J., Mood, A.M., Weinfeld, F. D., York, R. L. Equality of educational opportunity. Washington, D. C.: U.S. Government Printing Office, 1966.

Cooley, W. W. & Lohnes, P. R. Predicting development of young adults: Project TALENT five-year follow-up studies. (Interm report 5 to the U. S. Office of Education, Cooperative Research Project No. 3051). Palo Alto: American Institutes for Research and University of Pittsburgh, Project TALENT office, 1968.

Dunn, E. S., Jr. Review of proposal for a national data center. Statistical Evaluation Report No. 6. Washington, D.C.: Executive Office of the President, Bureau of the Budget, Office of Statistical Standards, December 1965. (Reprinted in United States House of Representatives, Committee on Government Operations, Subcommittee on Invasion of Privacy, The computer and invasion of privacy: Hearings before the Subcommittee. Washington, D. C.: United States Government Printing Office, 1966. Pp. 254-294).

Fano, R. M. & Corbató, F. J. Time-sharing on computers. Scientific American, 1966, 215 (3), 128-140

Flanagan, J. C., Cooley, W. W., Lohnes, P. R., Schoenfeldt, L. F., Holdeman, R. W., Combs, Janet, and Becker, Susan. Project TALENT one-year follow-up studies. (Final report to the U. S. Office of Education, Cooperative Research Project No. 2333.) Pittsburgh: University of Pittsburgh, Project TALENT Office, 1966.

Flanagan, J. C., Dailey, J. T., Shaycoft, Marion F., Orr, D. B, and Goldberg, I. Design for a study of American youth. Boston: Houghton Mifflin, 1962.

Flanagan, J. C., Davis, F. B., Dailey, J. T., Shaycoft, Marion F., Orr, D. B., Goldberg, I., and Neyman, C. A., Jr. The American high school student. (Technical report to the U. S. Office of Education, Cooperative Research Project No. 635.) Washington, D. C.: University of Pittsburgh, Project TALENT Office, 1964.

Glaser, W., and Bisco, R. Plans for the council of social science data archives. Social Science Information, 1966, 5 (4), 71-96.

Green, B. F., Jr. Digital computers in research. New York: McGraw-Hill, 1963.

Horst, P. Factor analysis of data matrices. New York: Holt, Rinehart & Winston, 1965.

Kaysen, C. Report of the task force on the storage and of access to government statistics. Washington, D. C.: Executive Office of the President, Bureau of the Budget, October 1966. (Reprinted in United States Senate, Committee on the Judiciary, Subcommittee on Administrative Practice and Procedure, Computer privacy: Hearings before the subcommittee. Washington, D. C.: United States Government Printing Office, 1967. Pp. 25-37.)

Kemeny, J. G. & Kurtz, T. E. Dartmouth time-sharing. Science, 1968, 162, 223-228.

Lohnes, P. R. & Cooley, W. W. Introduction to statistical procedures with computer exercises. New York: Wiley, 1968.

McCarthy, J. Information. Scientific American, 1966, 215 (3), 64-73.

Nasatir, D. Social science data libraries. The American Sociologist, 1967, 2, 207-212.

Sawyer, J. & Schechter, H. Computers, privacy, and the national data center: the responsibility of social scientists. American Psychologist, 1968, 23, 810-818.

Schoenfeldt, L. F. A national data resource for behavioral, social, and educational research. Palo Alto: American Institutes for Research, 1968.

Science Advisory Committee. Computers in higher education. Washington, D.C.: U.S. Government Printing office, 1967.

Shaycoft, Marion F. The high school years: Growth in cognitive skills. (Interim report 3 to the U. S. Office of Education, Cooperative Research Project No. 3051.) Pittsburgh: American Institutes for Research and University of Pittsburgh, Project TALENT Office, 1967.

Time Magazine, "Thinking about October." July 5, 1968, 92, 13.

## Table 1

## Project TALENT Follow-up Questionnaire

| Timetable | | | | |
|---|---|---|---|---|
| Grade in 1960 | 1-year Follow-up | 5-year Follow-up | 10-year Follow-up | 20-year Follow-up |
| 12th | 1961 (mailed) | 1965 (mailed) | 1970 | 1980 |
| 11th | 1962 (mailed) | 1966 (mailed) | 1971 | 1981 |
| 10th | 1963 (mailed) | 1967 (mailed) | 1972 | 1982 |
| 9th | 1964 (mailed) | 1968 (mailed) | 1973 | 1983 |

Table 2

MAJOR DATA ARCHIVES [a/]

| Name of Data Archive, Director | Address | Type of Data and Subject Matter |
|---|---|---|
| **Psychology and Education** | | |
| American Council on Education<br>Office of Research<br>Alexander Astin | 1785 Massachusetts Ave., N.W.<br>Washington, D.C. 20036 | Statistical data for American colleges and universities including school characteristics; national norms for the freshman class; e.g. 359 institutions and 280,650 students in 1967. |
| Management Research Center<br>University of Rochester<br>B. M. Bass | College of Bus. Admin.<br>1409 Temple Building<br>Rochester, New York 14604 | Managers' responses from over 30 countries on 12 simulations of key management activities; values and behavior of managers in various organizations and cultures. |
| National Education Association | 1201 16th Street, N.W.<br>Washington, D.C. 20036 | Three data banks--one on staff-school board negotiation agreements; two with demographic and financial data, one on all states and territories since 1949, and one on large school systems with enrollments of 25,000 or more. |
| New England Education Data Systems<br>Michael J. Wilson | 222 Alewife Brook Parkway<br>Cambridge, Mass. 02138 | Master file of information on students in participating New England schools with descriptive data on the schools, accumulated grades by subject and teacher. Local and national norms for test results. |
| Project TALENT Data Bank<br>Lyle F. Schoenfeldt | P.O. Box 1113<br>Palo Alto, Calif. 94302 | Probability sample of 440,000 high school students: abilities, interests, plans, activities and background. Follow-up data collected 1 and 5 years after h.s. graduation. |
| **Government** | | |
| Bureau of the Census<br>A. Ross Eckler | Washington, D.C. 20233 | Census data; population estimates; migration statistics; business and trade figures updated monthly. |
| Bureau of Labor Statistics<br>Arthur M. Ross | U.S. Department of Labor<br>Washington, D.C. 20212 | Data on manpower and employment; wages and industrial relations; productivity, technology and growth; industrial safety; prices and living conditions. State and area data. |
| **Social Sciences: Sociology, Economics and Political Science** | | |
| Roper Public Opinion Research Center<br>Philip K. Hastings | Williams College<br>Williamstown, Mass. 01267 | Sample surveys from 26 American sources and 43 foreign countries covering politics, economics and business, education, health and welfare, occupations, public opinion, organizations, community problems, intergroup relations. |
| Inter-University Consortium for Political Research<br>Warren E. Miller | University of Michigan<br>P.O. Box 1248<br>Ann Arbor, Michigan 48106 | Partially integrated collections of survey data, aggregative statistics and public record data from the early 1800's. Data covers ethnic origins, religious affiliations, status, location and mobility, political partisanship and socialization, occupational and income distributions and attitudes. |
| International Data Library and Reference Service<br>David Nasatir | Survey Research Center<br>2220 Piedmont Avenue<br>University of California<br>Berkeley, California 94720 | Sample survey materials covering politics, communications, and many other aspects of social behavior and public opinion. Special emphasis on Asia and Latin America. |
| National Opinion Research Center<br>Norman Bradburn | University of Chicago<br>6030 South Ellis Avenue<br>Chicago, Illinois 60637 | Sample surveys on U.S. and regions of U.S., covering health and welfare, occupations and professions, mass communications, economics and business, politics, education, community problems, intergroup relations, methodology and theory. |
| Bureau of Applied Social Research<br>Alan H. Barton | Columbia University<br>605 West 115th Street<br>New York, New York 10025 | Sample surveys including some panel studies, some collections of social statistics. Data for particular groups, mostly in U.S., on health and welfare, occupations and professions, mass communications, politics, education, organizations. |
| School of Public Health and Administrative Medicine Research Archives<br>Jack Elinson<br>Beatrice Mintz | Columbia University<br>630 West 168th Street<br>New York, New York 10032 | Sample surveys and operational data from sociometrical research on administrative medicine and public health. Areas: New York State, New York City, other U.S. and Puerto Rico. |
| International Development Data Bank<br>Jack Bain | Michigan State University<br>426 S. Kedsie Hall<br>East Lansing, Mich. 48823 | Sample surveys conducted in less developed countries, primarily Latin America, Africa and Asia--social science studies of development and modernization. |
| Council for Inter-Societal Studies<br>Richard D. Schwartz | Northwestern University<br>1818 Sheridan Road<br>Evanston, Illinois 60201 | Journal abstracts, microfilm texts, some sample surveys and legislative data for Indiana, 1939-1961; texts on political parties in Japan, Nigeria, Argentina, Norway, Yugoslavia. |

Table 2 continued

MAJOR DATA ARCHIVES

| Name of Data Archive, Director | Address | Type of Data and Subject Matter |
|---|---|---|
| **Social Sciences (continued)** | | |
| Human Relations Area Files<br>Clellan S. Ford<br>Frank W. Moore | Yale University<br>P.O. Box 2054 Yale Station<br>New Haven, Conn. 06520 | Texts on social structure and culture of over 250 societies throughout the world organized by ethnic and political units--dietary practices, kinship, social organization. |
| Yale Economic Growth Center<br>Gustav Ranis | Yale University<br>52 Hillhouse Avenue<br>New Haven, Conn. 06520 | Analyses for 20 less developed and 15 industrial countries. |
| UCLA Political Behavior Archive<br>Dwaine Marvick | Dept. of Political Science<br>University of California<br>Los Angeles, Calif. 90024 | Opinion surveys, mail questionnaires, electoral and demographic statistics on politics, elite behavior, electoral phenomena and various aspects of public opinion and campaign activity. Emphasis on Southern California with some foreign data. |
| Latin American Data Bank<br>Thomas L. Page | Dept. of Political Science<br>University of Florida<br>Gainesville, Fla. 32601 | Population, agricultural and industrial census detail cards, election returns, economic time series; aggregate social, economic and political data on Costa Rica, Guatemala, Columbia. |
| Survey Research Laboratory<br>William Flecke | 3 David Kinley Hall<br>Univerisity of Illinois<br>Urbana, Illinois 61810 | Sample surveys, economic and voting statistics, politics, some aspects of public opinion and social behavior. Area: Illinois and nearby states |
| Laboratory for Political Research<br>George R. Boynton | Dept. of Political Science<br>University of Iowa<br>Iowa City, Iowa 52240 | Surveys in Iowa from polls; studies of local party leaders in Massachusetts and North Carolina. Socioeconomic and career patterns of Argentine politicians. |
| Public Opinion Survey Unit<br>David Lenthold | Research Center, School of Business & Public Administ.<br>University of Missouri<br>Columbia, Missouri 65201 | Sample surveys of the public and of elites on politics and public opinion; some election statistics, studies of political biographies, and the census. Area: American states, particularly Missouri. |
| Louis Harris Political Data Center<br>Alden Lind | Dept. of Political Science<br>University of North Carolina<br>Chapel Hill, N.C. 27514 | Public opinion surveys collected since 1956 on policies in individual U.S. states; attitude research. |
| Social Science Training and Research Laboratory<br>William T. Liu | University of Notre Dame<br>Notre Dame, Indiana 46556 | Political, sociological, and population data. Also, Colonial Economic History Data Archive on colonial customs and shipping. |
| Archive on Comparative Political Elites<br>Lester Seligman<br>Arthur M. Hanhardt, Jr. | Dept. of Political Science<br>University of Oregon<br>Eugene, Oregon 97403 | Comparative data for United States, Israel, East and West Germany, and the developing nations on recruitment, mobility, perceptions and background of political elites. |
| Social Science Data and Program Library Service<br>Mrs. Margaret M. Adams | Social Systems Research Inst.<br>Rm 4451 Social Science Bldg.<br>University of Wisconsin<br>Madison, Wisconsin 53703 | Survey of consumer finances; business loan surveys of member banks of Federal Reserve; industrial data and other census data. Areas: American cities, some files for Wisconsin alone. |
| **Foreign** | | |
| Carleton University Social Science Data Archives<br>Roman R. March | Dept. of Political Science<br>Carleton University<br>Colonel By Drive<br>Ottawa 1, Canada | Sample surveys on politics, public opinion and other social behavior; biographies; election statistics; census data. Area: Canada. |
| Carnegie Human Resources Data Bank<br>H. Howard Russell | The Ontario Institute for Studies in Education<br>102 Bloor Street West<br>Toronto 5, Ontario, Canada | Test and questionnaire data from high school students as they progress through high school, grades 9 through 12: Academic aptitude and achievement tests, Canadian Interest Inventory, school marks and student and teacher questionnaires |
| Ontario Statistical Center<br>Kenneth Cheng | Ontario Treasury Department<br>950 Younge Street<br>Toronto 5, Ontario, Canada | Economic time series data bank--monthly, quarterly and annual data carried back where possible to 1946. Population statistics by Provinces; immigration statistics; income, expenditure and production data; vital statistics. |
| Social Science Research Council Data Bank<br>A. M. Potter | University of Essex<br>Wivenhoe Park<br>Colchester, Essex, England | Social surveys sponsored by government departments, commercial organizations and academic researchers. |
| Steinmetz Institute | Herengracht 457<br>Amsterdam-C.<br>The Netherlands | 100 studies from around the world on aspects of social behavior such as birth rate, family planning, emigration, religion, housing. Public opinion studies on political events and standards of living. |
| Zentralarchive<br>Ervin K. Scheuck | Zentralarchive for Empirical Social Research<br>University of Cologne<br>5 Cologne, West Germany | Surveys from Germany and German language area. Data on politics and ideology, economics and business. |

a/ Adapted from Nasatir (1967).